DOCUMENT RESUME

ED 381 168 IR 055 461

TITLE Information and Networking Technologies in Russian Libraries. UDT Occasional Paper #1.

INSTITUTION International Federation of Library Associations and Institutions, Ottawa (Ontario). International Office for Universal Dataflow & Telecommunications.

PUB DATE 95

NOTE 24p.

PUB TYPE Collected Works - General (020)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Access to Information; *Character Recognition; Computer Networks; Foreign Countries; Higher Education; *Information Technology; *Library Networks; Online Catalogs; *Online Systems; *Optical Scanners; Telecommunications

IDENTIFIERS Internet; *Russia; Universal Dataflow and Telecommunications

ABSTRACT

The Universal Dataflow and Telecommunications (UDT) Occasional Papers distribute information on the use of networking, information technology and telecommunications by and of interest to the international library community. This occasional paper is comprised of three papers related to technologies in Russian libraries: (1) "The First Russian Computerized Library Network: Description and Perspectives of the LIBNET Project" (Yakov Shraiberg and Mikhail Goncharov); (2) "The Current State and Prospects of Online Systems in Russian Libraries" (Yakov Shraiberg); and (3) "Problems of Optical Character Recognition Technologies in Russian Libraries and Information Centres" (M. Goncharov; and D. Nikolaev). (MAS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

International Federation of Library Associations and Institutions

Universal Dataflow and Telecommunications

ED 381 168

Information and Networking Technologies in Russian Libraries

1995

UDT Occasional Paper #1

IR 055 461

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

A.L. Van Wesemael

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

UDT Occasional Paper #1: Information and Networking Technologies in Russian Libraries: April 1995

Contents

1. Shraiberg, Yakov, The First Russian Computerized Library Network - a Description and Perspectives of the LIBNET Project.

2. Shraiberg, Yakov, The Current State and Prospects of Online Systems in Russian Libraries.

3. Goncharev, M. and D. Nikolaev, Problems of Optical Character Recognition in Russian Libraries and Information Centres.

UDT Occasional Papers

The UDT Occasional Papers allow the UDT Core Programme to distribute information on the use of networking, information technology and telecommunications by and of interest to the international library community.

They will be published as received in their original language and will also be available on the IFLANET WWW server.

The First Russian Computerized Library Network: Description and Perspectives of the LIBNET Project

Dr. Yakov Shraiberg and Mr. Mikhail Goncharov
The Russian National Public Library for Science and Technology
Moscow, Russia

In this decade we have become witnesses of an intensive introduction of computerized systems into Russian libraries. A system of library networks has been formed in the country, though not in the telecommunication but in the organizational and functional sense. Every network has its own leading library (or libraries). As a rule, these are large scientific libraries with impressive collections and well-developed computerized systems. They are either state or branch depositaries responsible for a methodological and consultative guidance provided to other libraries of the network. The Russian National Public Library for Science and Technology, for instance, is the leading library for a network of scientific and technological libraries of the country, the State Russian Library (the former Lenin Library) and the Russian National Library (the former Saltykov-Shchedrin Library) are both the leading ones for the Ministry of Culture libraries, the State Central Scientific Medical Library is the leading one for a network of medical libraries, etc. Interlibrary cooperation has always been in force incorporating ILL services, acquisition coordination, joint updating of the Union Catalog, and information retrieval on query. However, it has been based on traditional technologies. Introduction of the RELCOM e-mail network, a home analog of INTERNET, has drastically changed the situation, though leaving unsolved the tasks of interactive access to electronic catalogs, to the Union Catalog and the National Bibliography System, a query-response mode, and other real-time modes. That is why, a creation of the first Russian computerized library network has become the order of the day.

Thus, a project of the State Federal Library Network was included by the State Council for Informatization into the Informatization Concept of Russia. After that, an interbranch commission was set up in 1992 under the aegis of the Russian Ministry of Culture to develop and introduce a project of a computerized library information network. Experts of six major Moscow libraries formed a development group and, at the end of the same year, submitted a project of the Russian Federal Library Information Network. It was called LIBNET, which is an acronym of the LIBrary NETwork.

The first information on the project was published in [1]. It read that LIBNET would:

- give further impetus to the development of library cooperation and integration on the basis of modern hardware and software,
- considerably update the respective library technologies and improve user service, including real-time access of users to the collections of the participating libraries,
- provide access to foreign networks, data banks, and electronic catalogs of the largest world libraries.

The key idea of the project was to create a ramified network incorporating library information networks, separate computers of Russian libraries, and individual users.

The concept was based on the utilization of:

- a unified LAN environment,
- UNIMARC format that is a communicative environment operating with various types of formats selected by each participating library for computerization of its local technologies,
- a unified distributed electronic catalog of the collections of the participating libraries.

The concept led to a pilot project developed by the Russian National Public Library for Science and Technology and the Institute for Information Transmission Problems in collaboration with the Moscow University Scientific Library and the State Central Scientific Medical Library.

At stage 1 LIBNET was to provide users with online access to electronic catalogs and databases of six major Moscow libraries, including access to the joint product - the Russian Union Catalog of Sci-Tech Publications maintained by the Russian National Public Library for Science and Technology.

The six libraries joining the project at stage 1 were:

- the Russian National Public Library for Science and Technology,
- the Moscow University Scientific Library,
- the State Central Scientific Medical Library,
- the Russian Academy of Sciences Natural Sciences Library,
- the Central Scientific Library of Agriculture,
- the State Public Historical Library.

At the beginning of 1993, at the start of the project's implementation, computer catalogs of the Russian National Public Library for Science and Technology had about one million records. The Library maintained a double-mainframe computer system and NOVELL PC-based LAN. Telecommunication access was provided to the electronic catalog via switched channels, information products could be accessed via X.25 (the Russian "ROSNET" network, international address: 02506100240), and remote order of publications was in operation.

The other participating libraries had electronic catalogs of different sizes. The largest one was in the State Central Scientific Medical Library and had 300,000 records.

Only three libraries, namely, the Russian National Public Library for Science and Technology, the State Central Scientific Medical Library, and the Moscow University Scientific Library, had local NOVELL networks. The rest ones worked with local computerized systems.

That is why, four tasks were recognized as primary at stage 1:

- to install local NOVELL networks in three libraries (the Russian Academy of Sciences Natural Sciences Library, the Central Scientific Library of Agriculture, and the State Public Historical Library) and ensure the operation of their computerized systems in the network environment,
- to generate the NETWARE ACCESS SERVER package in each LAN,
- to examine public telephone channels and select a portion with the highest reception/transmission specifications to be further used as the basic communication environment,
- to purchase telecommunication equipment, install modems, purchase and test telecommunication software.

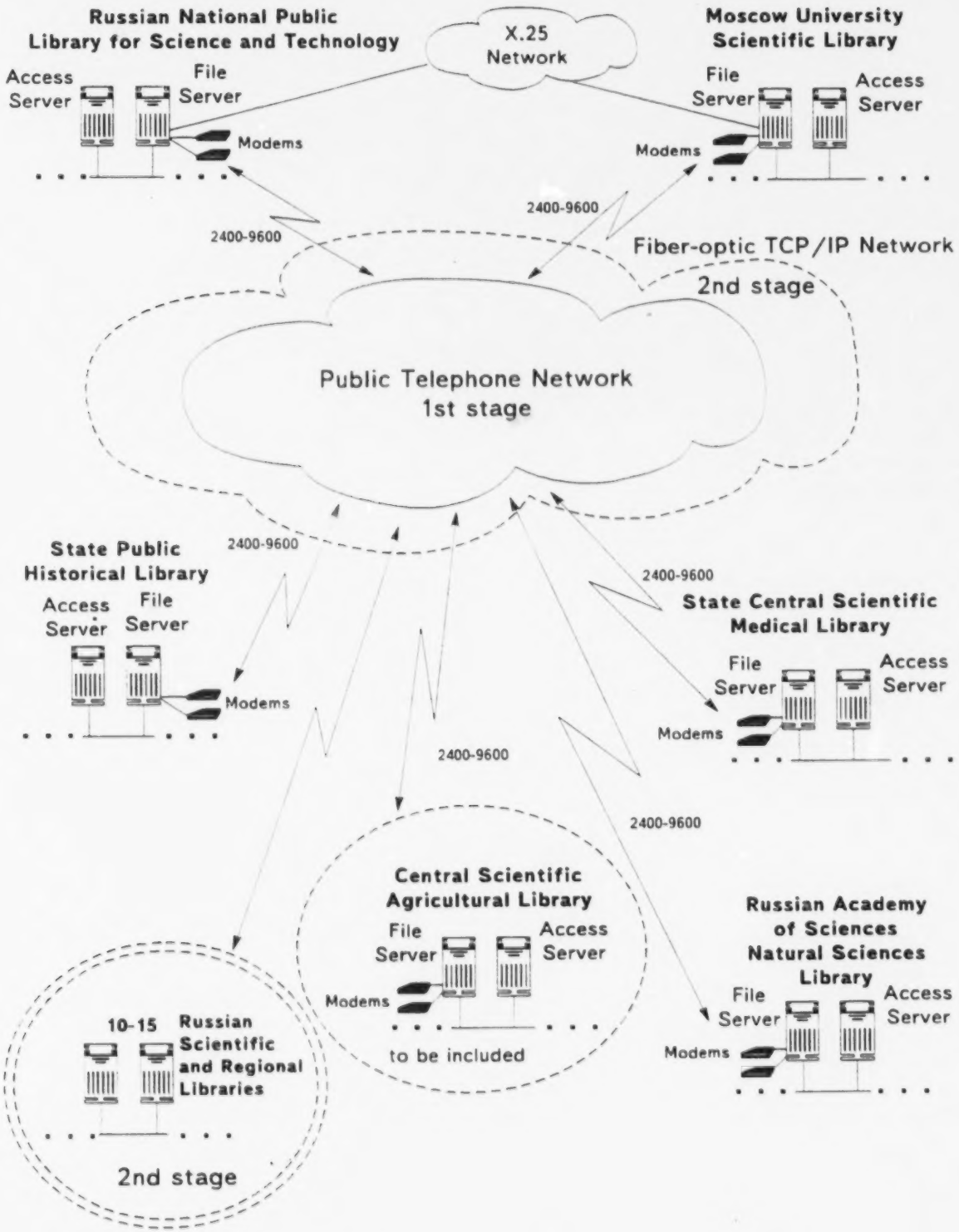


Fig 1. LIBNET Network

7

At stage 1 the group was joined by the Institute for Information Transmission Problems that was made responsible for telecommunication tasks. Meanwhile, it was decided that the Central Scientific Library for Agriculture would join the project at stage 2 due to its technological lag.

Stage 1 of the LIBNET project was completed to the June of 1993. Since then, five major Moscow libraries have gained access to the electronic catalogs of each other, over 500,000 records reflecting the electronic catalogs of the participaring libraries became available for many remote users, and every user equipped with an IBM PC or an ANSI terminal, a HAYES-compatible modem, and a telephone can access electronic catalogs of these libraries, search for and request publications, and create personal information files.

LIBNET is the first Russian special library network and its introduction was a true success of Russian libraries (Fig.1).

On stage 2 LIBNET will develop along the following lines:
Connection of other Russian and Moscow libraries, so that their number will reach 10 - 15 at the end of this year. Resources of still another large Russian libraries and information centres with well-developed systems of electronic catalogs and databases are planned to be connected to LIBNET at the following stages of the network development. Note should be made that Russian provincial libraries are interested in joining the project as they realize its advantages in distributed bibliographic processing and access to large data banks and Union Catalogs. At the same time, Russian provincial libraries possess unique information on local lore, history, economy, ecology of specific regions, and other data sometimes unavailable in the collections of central libraries.

1. Transition from switched channels to X.25 and TCP/IP communication environments, including fiber-optic and packet-radio channels. Some experience has already been accumulated in this field, as the Russian National Public Library for Science and Technology and the Moscow University Scientific Library are already connected to X.25.
   An intermediate and, later on, an alternative stage at this point will be a connection of libraries to the channels of "ISKRA-2" - a specialized home-made fault-tolerant network.

2. Creation of an interlibrary network supporting electronic catalogs, the National Bibliography System, ILL, and other library functions. Within LIBNET, the libraries will maintain an integrated database for coordinated foreign acquisition (the Russian National Public Library for Science and Technology will be made responsible for database maintenance). The libraries will adhere to the principles of a unified cataloging of foreign publications acquired via online and E-mail access to foreign bibliographic databases or foreign bibliographic CD-ROMs according to the UNIMARC format.

In conclusion, we would like to repeat that LIBNET will considerably improve the quality of services offered to scientists, engineers, students, and all information users in the

Moscow region and Russia in general, push up library technologies and ensure the creation of the library computer network of the whole country.

The project will ensure the establishment of modern network infrastructure for information business in Russia and allow Russian holders of information resources to take part in international cooperation in this field, which is sure to meet the requirements of Russian and foreign users.

And, finally, one of the most important and vitally necessary perspectives is mutual online access to and information sharing with foreign libraries and information centres via facilities of the global INTERNET.

Russia and many countries of the world have been looking forward to this perspective and now it is near at hand.

Reference

1. Shraiberg, Yakov L., Goncharov, Michail, V. "LIBNET - Project of the Major Russian Libraries Network". Abstracts of the 15th Biennial IATUL Conference (July 1993), Technical University Hamburg-Harburg, D-5.

The Current State and Prospects of Online Systems in Russian Libraries

Dr. Yakov Shraiberg
First Deputy Director
The Russian National Public Library for Science and Tec'..ology
Moscow, Russia

A new era in automation of Russian libraries that began in the 1990s has manifested itself in intensive introduction of personal computers and LANs into library practice. In the era of large mainframe computers of the IBM/360/370 type and their analogs, only major libraries could automate their technologies as they alone had enough money for these expensive and powerful computers, while their vast collections and processing flows could ensure more or less effective operation of the mainframes.

IBM/360/370-compatible mainframes were installed in several Russian libraries:

1. The Russian National Public Library for Science and Technology. It had over one mln. bibliographic and address-reference machine-readable records run on a two-computer mainframe complex, an IBM/370 - compatible computer, that was a joint product of several COMECON countries. The most part of the mainframe storage was occupied by the databases of the Union Catalog of Sci-Tech Publications acquired by the FSU libraries.

2. The Russian Academy of Sciences Natural Sciences Library. It had two mainframes storing thousands of bibliographic records on domestic and foreign publications, over forty thousand reference records on library users, and several hundred thousand records on ILL requests.

3. The State Public Library for Science and Technology of the Russian Academy of Sciences Siberian Branch. It had large information amounts generated by major Russian STI centres, like the All-Russian Institute of Sci-Tech Information (abstracts), All-Russian Sci-Tech Information Centre (R&D reports and dissertations), and "Poisk" Research and Production Association (patents). This library had several mainframes belonging to the Multi-user Computer Centre of the Russian Academy of Sciences Siberian Branch and used mainly for information services, instead of library automation.

Mainframes were also used by several other libraries, such as the Central Scientific Library of Agriculture, the All-Russian Technological Patent Library, and many university libraries which rented them in the university computer centres.

While a few tasks of library automation and, to a greater extent, information tasks were solved due to the application of mainframes, online technologies were very much under-developed. There existed a V21 (300 baud) switched access by means of a teleprocessing processor, but data exchange and access to foreign databases were provided sporadically and to a different degree of success via VNIIPAS (The Union

Scientific and Research Institute for the Problems of Automated Systems) that worked as a mediator (and, unfortunately, as a "bottle-neck" for the tasks of external teleaccess).

The situation has changed drastically over the last three years. Many IBM PC-compatible computers have been imported, and a competition with them has turned senseless for home-made PCs - EC-1840 series computers, ISKRA, NEURON, and the like. Russian libraries are equipped now with IBM PC/AT-286,386,486 and even Pentium computers with Gb hard disks, SCSI-adapters, and other modern characteristics, PC-based LANs, CD-ROM databases, scanners, desk-top publishing systems, and many other applications of modern information technologies.

PC-based LANs have been introduced by virtually all Russian libraries. They have become a basis for development and maintenance of local automated systems. Many libraries have initiated a gradual development of their systems using CDS/ISIS/M, Clipper, Paradox, and other packages, C, C+, C++, and Pascal languages. But major libraries are also actively developing their own integrated systems and supplying them to libraries that do not have such an opportunity. Foreign products for library automation have not become popular in Russian libraries because of a permanent money problem for their purchase (the only exception in this respect is a French "LIBER" system, thanks to its cheapness). In telecommunication, our success is very modest, but every single day brings some progress. The results are especially encouraging in data exchange via e-mail. RELCOM network as the Russian node of the global INTERNET has linked many libraries and provided an opportunity for exchanging information and messages, generating technologies for ILL services and multi-user databanks, and offering remote access in the pending-request mode..

We shall try to bring our information in line with the types of libraries, though it is not easy, in some cases, to define the type according to the accepted library classification: comprehensive, special, public, university, etc. Very often we shall deviate from this model with regard to the Russian national features (Table 1).

Table 1

N	Types	Major Libraries	Subordination	Number of Libraries	Total Collection (items)	Users
1.	Sci-Tech Libraries	The National Public Library for Science and Technology (Moscow)	The Ministry of Science and Technological Policy Branch Ministries	7,000	20 mln	5 mln
2.	Academy of Sciences Libraries	The Academy of Sciences Library (St.Petersburg) The Academy of Sciences Natural Sciences Library (Moscow) The State Public Library for Science and Technology of the Academy of Sciences Siberian Branch The Institute for Information on Social Sciences	The Russian Academy of Sciences	370	55 mln	200,000
3.	Universal, Public, Municipal Libraries	The Russian State Library (Moscow) The Russian National Library (St.Petersburg)	The Ministry of Culture	50,000	900 mln	52 mln
4.	University and College Libraries	The Moscow University Scientific Library (conditional supervision)	The Ministry of Higher Education	700	300 mln	4 mln
5.	Medical Libraries	The State Central Scientific Medical Library	The Ministry of Public Health	1,500	50 mln	2,5 mln
6.	Agricultural Libraries	The Central Scientific Library of Agriculture	The Academy of Agricultural Sciences	600	40 mln	1,5 mln
7.	Others			10,000	200 mln	10 mln

Below we shall examine the state-of-the-art online access in the major libraries of every type. Note, that the primary technology that is used today for online access to the resources of Russian libraries is the LIBNET network, the project of which and application results are described in this bulletin. Therefore, we shall refer to LIBNET, when necessary, instead of brouching the respective online modes.

1. The System of Sci-Tech Libraries

This is a three-layer system with its centre in the Russian National Public Library for Science and Technology, one of the major Russian libraries. The Library is the state depository of sci-tech literature and a holder of the Union Catalog.

In general, the System of Sci-Tech Libraries comprises 7,000 libraries with the collections carrying 20 mln items.

The second, branch, layer of the System comprises 17 central Sci-Tech libraries of the Ministries and Departments. These libraries are, in turn, considered as central for the corresponding departmental library systems operating in branch institutions and enterprises (the third layer). All central Sci-Tech libraries and 10 percent of the third-layer libraries have passed the initial stage of automation and are equipped with IBM PC-type computers. Some libraries still utilize mainframe computers of institutions by the right of a share-user.

The most automated library in the country today is the Russian National Public Library for Science and Technology. It has the most developed telecommunication technology, compared to other Russian libraries. The Library's online and telecommunication system is shown in Fig.1.

The Library's telecommunication access is based on a special access server allowing for sixteen concurrent sessions. NetWare Connect 1.0 and NetWare Access server.1.3 are the principal software packages. Presently, telecommunication is provided to over 25 databases, the Library's electronic catalog, electronic publications, and address-reference products via four switched telephone lines of the public system telephone network (PSTN) and eight logical channels by means of a packet-switched X.25 network. The address of the telecommunication system is as follows: 2506100240. Besides, INTERNET provides access in the pending-request mode. Its address is: POISK@gpntb.msk.su.

Online access via INETRNET will be put into operation within the coming two or three months. The Library offers concurrent access to the INTERNET resources from four user workstations.

CD-ROM Networking is a new trend in the Library's activities. A special network based on the CD-ROM-server has already been developed and launched in the Library. It allows for multi-access to the Library's CD-ROM databases of about 100 concurrent users.

The Library's online and telecommunication system

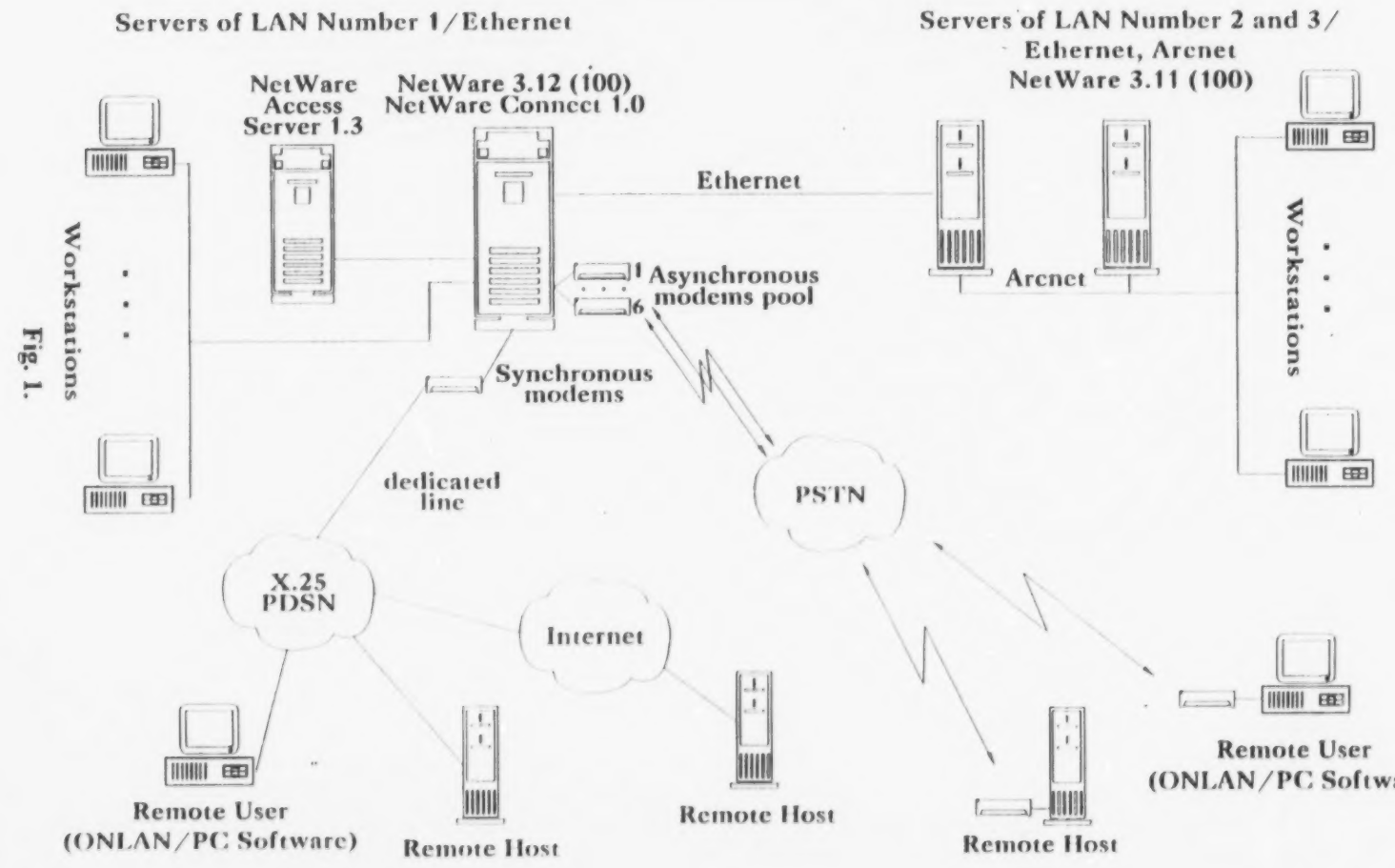


Fig. 1.

2. Information Library Network of the Russian Academy of Sciences

The Network consists of 2 regional and 2 branch systems. The regional centres are:

1. The Russian State Public Library for Science and Technology of the Russian Academy of Sciences Siberian Branch.

2. The Academy of Sciences Library (St. Petersburg).

The first one unites 67 libraries, the second one - 41 libraries.

The branch centres are:

1. The Russian Academy of Sciences Natural Sciences Library.

2. The Institute of Information on Social Sciences (the system of humanitarian and social orientation).

The first one unites 246 libraries and the second one - 22 libraries.

Library holdings of the system total 55 mln. items. Automation in these libraries began in the late 1970s and the Natural Sciences Library network has become the major one. Its "NAUKA" (Science) automated system provides automated information-library technologies for the central library and libraries of the Academy of Sciences research institutes.

In May 1994 the Library joined the LIBNET project and started to offer telecommunication access to a part of its collections run on eight personal computers linked into LAN. Teleaccess is provided via one switched PSTN line as yet. The system is developing within the LIBNET project that envisages the utilization of "ISKRA-2" specialized departmental fault-tolerant telephone network and fiber-optic INTERNET channels. A transition to fiber-optic channels this year is envisaged by the LIBNET development. Telecommunication access to the Library's resources is shown in Fig.2. This layout is taken as model by many Russian libraries.

3. The Russian Ministry of Culture Library System

It embraces some 50,000 universal, scientific, and public libraries with the holdings over 900 mln. items. The Russian State Library (the former Lenin Library) and the Russian National Library (the former M. E. Saltykov-Shchedrin Public library in St. Petersburg) are the leading libraries within this system. In 1986, the Ministry of Culture Computer Centre together with the specialists of the above two libraries started to automate their library functions. A project of a standard automated library-information system based on IBM PC LAN was developed. Its implementation began in 1993 in many libraries of the country.

The Russian State Library is the major library of this system and the country as a whole. Despite the automated system and a big park of computers (three HP-3000/48 minicomputers and some eighty IBM PCs), the Library does not have online access and has not joined the LIBNET project thanks to the absence of LAN. Note, that the basic idea of the LIBNET project is to form a unified global NOVELL network by uniting local NOVELL networks of the participating libraries. Direct online connection to INTERNET resources and participation in the LIBNET project are considered as the Library's prospects.

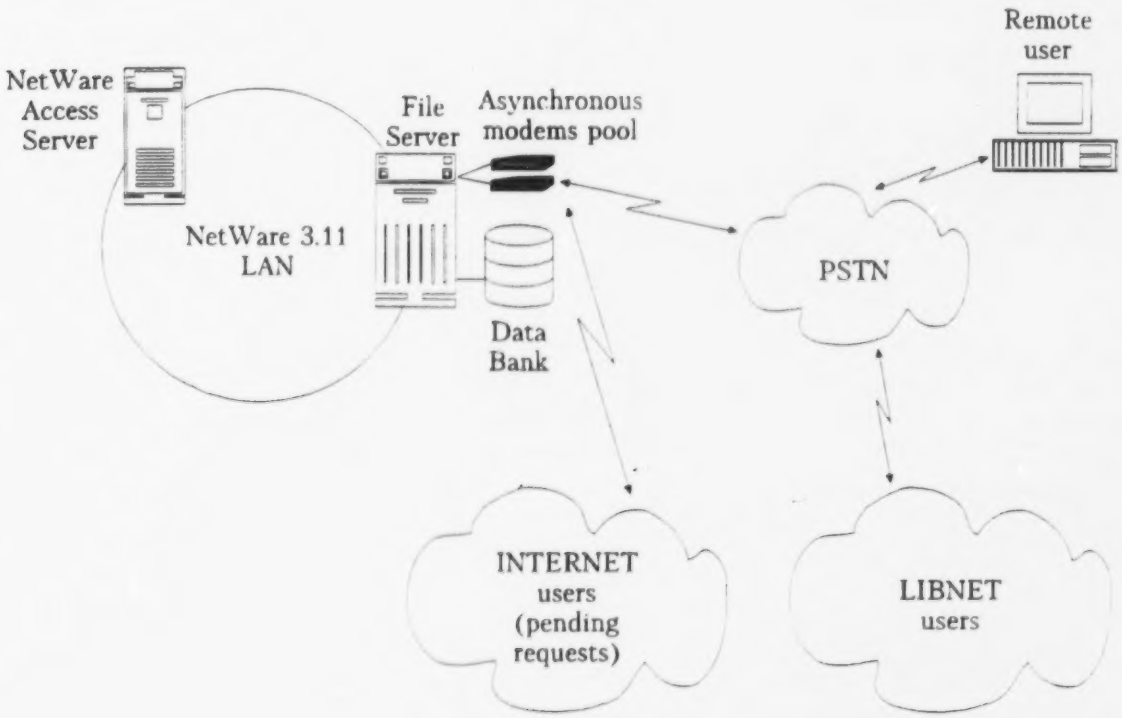


Fig. 2.

4. Russian University Library System

It unites more than 700 libraries with the collections of over 300 mln. units. The level of their automation is rather high, every second library is equipped with several IBM PC/AT computers and the total number of them is over 2,500. The total number of computer records of university libraries is over 2 mln, but many of them are duplicates. Most of these libraries are utilizing domestic software, either "Biblioteka" (developed by the Moscow University Scientific Library) or "MARK" system (developed by "Informsistema" Scientific and Production Association). There is also an automated system known as BKS, which was developed by the Moscow State Technological Library.

The "Library Computer Network" Joint-Stock Company, set up on the basis of the Moscow University Scientific Library by a special decree of the Russian Federation Committee on Higher Education, is coordinating the automation activities of university libraries. The first stage of the computer network of university libraries and regional networks in St. Petersburg and Chelyabinsk is to be launched this year together with the central catalog of the Moscow University Scientific Library (the catalog volume in 1994 - 100,000 records).

The Moscow University Scientific Library is the leading library of the network. It has sixty personal computers linked into LAN and a well-developed telecommunication network: PSTN - eight concurrent- access channels, X.25 PDSN and direct connection to INTERNET. Virtually, it is the only library in Russia with an IP address for online user's access to INTERNET. Moreover, the Library is a member of LIBNET and, together with the Russian National Public Library for Science and Technology, utilizes INTERNET e-mail for the work in the pending-request mode.

5. Branch Systems

Networks of medical and agricultural libraries are the largest ones among branch systems. The medical library network includes 1,500 libraries and is headed by the State Central Scientific Medical Library. It is one of the major special libraries in the country and functions within the system as a centre for cataloging, subject analysis, and information service. Its holdings carry over 2 mln. items. The Library has more than fifty personal computers linked into LAN. The Library joined LIBNET in 1993 and, since then, has been providing access to its electronic catalog carrying over 300,000 bibliographic records. Telecommunication has been introduced in accordance with the layout shown in Fig.2. The utilization of fiber-optic channels of LIBNET and "ISKRA-2" networks is the Library's prospect.

Agricultural library network comprises 593 university libraries, technical secondary schools, and experimental stations and institutions.The total collection of the agricultural network amounts to 40 mln. items.

The Central Scientific Library of Agriculture is the leading library of the network with its departments in St. Petersburg and Novosibirsk. The Library has twenty IBM PC-compatible computers linked into NOVELL LAN, but online access is unavailable as yet. The Library is being connected to LIBNET and, in a month, will have telecommunication system corresponding to that in Fig.2. All telecommunication prospects are connected with the application of the LIBNET network.

6. Other Types

There are other library types in the country: the Parliamentary Library, the President Administration Library, systems of theatre, trade union, school libraries, and the like.

The President Administration Library is the only one to actively introduce automation. The Parliamentary Library has more than twenty IBM PC/386/486 computers but is just beginning its automation under CDS/ISIS/M and TINLIB (UK). Its online access systems are under development and will utilize both INTERNET and LIBNET resources.

Other libraries, not pertaining to any of the above systems, have no telecommunication, but many of them are users of the Russian National Public Library for Science and Technology and some other libraries of the country.

Conclusions

1. Development of online access in Russian libraries is at the initial stage, though several libraries have made a breakthrough.

2. LIBNET gave a great impetus to the development of telecommunication systems in Russian libraries. At stage 1 it united six Moscow libraries and is planning to add some ten-fifteen more (with several hundred users) by the end of this year.

3. INTERNET is being implemented intensively. Most of the major and medium Russian libraries have e-mail, and several libraries are working in the pending-request mode via e-mail. Some of them have started the application of online INTERNET facilities and fiber-optic channels.

4. Several libraries are developing other telecommunication access systems, for instance, via X.25 PDSN.In addition, the leading libraries in automation and telecommunication are developing new projects. As an example, we may take BIBNET initiated by the Russian National Public Library for Science and Technology and the Moscow University Scientific Library. BIBNET (Fig.3) is a project for communication via X.25 dedicated channels and TCP/IP in the single-vendor UNIX environment.

References

1. Shraiberg, Yakov. The State of Library Automation in Russia. Microcomputers for Information Management, 10(4), 1993, 293-310.

2. Shraiberg, Yakov. Automated Systems of Russian Libraries and their Online and CD-ROM Service. Proceedings of the International Seminar on Information Technologies and Information Services. Shanghai, China, October 20-24, 1994, 181-186.

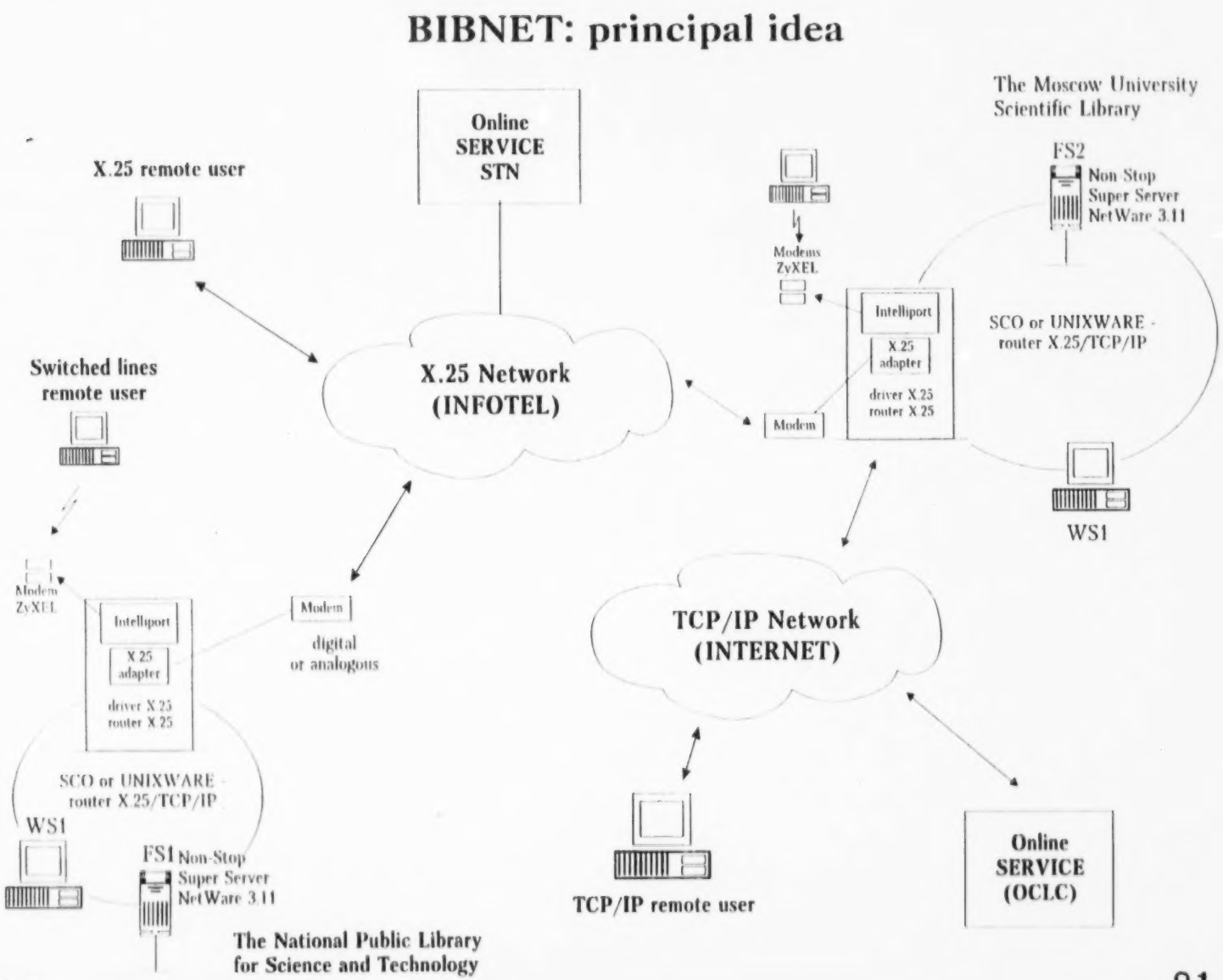


Figure. 3.

BEST COPY AVAILABLE

Problems of Optical Character Recognition Technologies in Russian Libraries and Information Centres

M. Goncharov, Department Head
D. Nikolaev, Section Head
The Russian National Public Library for Science and Technology
Moscow, Russia

Development and application of new processing technologies for large information amounts on various media is in the limelight of attention in Russia. The importance of this task for Russian libraries and information centres is even greater than for their Western partners due to late introduction of modern automated library-information systems and bibliographic information entry in a machine-readable form.

Large libraries that have already introduced such systems into the local technologies have met with a problem of retrospective conversion of information from catalogs of many millions cards and formation of full-text and hyper-text databases of unpublished editions and "grey" literature.

The problem is twofold: (a) software and technological, (b) economic.

In this paper, we shall describe briefly the first part of the problem - application of optical character recognition systems in Russian libraries and information centres.

1. In the late 1980s the leading libraries and information centres of the country, namely The Russian National Public Library for Science and Technology, The Moscow University Scientific Library, and The All-Russian Centre for Sci-Tech Information, started the reseach with the purpose to create automated systems for information entry from hard media. However, the research has yielded almost no significant results due to the absence of reliable software packages for the Cyrillic alphabet recognition. Note, that the algorithms for the Cyrillic alphabet recognition and their usage differ from those for the Latin alphabet. Unlinking characters minimize an opportunity to utilize even the most advanced and expensive Western programs for cyrillic texts. Letters "Û" and "É" cannot be duely processed and characters "Ô" and "Æ" are almost unrecognizable either. The impossibility to use spell-check for cyrillic texts turned out as still another problem, while spell-check is an inseparable part of all character recognition technologies.

Several Russian companies have tried to create new software products, but, due to the complexity of the problem, positive results have been achieved only recently.

With creation of a comparatively reliable recognition software many Russian libraries and information centres started its intensive introduction into local processing technologies for hard media. On this way, however, they met with still another problem hindering the application of software, i.e. low quality of original texts.

While the recognition accuracy for publications from western journals of a high polygraphic level is about 98 percent, it is much lower for publications produced on domestic polygraphic basis as some characters may stick together in a word or remain unprinted. Moreover, the accuracy of recognition is lowered due to the quality of cards printed on home-made typewriters as, constituting the most part of library card catalogs, they are not void of the above drawbacks.

The most developed is scanning of journals and newspapers, since, in most cases, their polygraphic quality corresponds to the Western standards.

As for the most significant tasks, such as retrospective conversion and creation of electronic depositories of domestic publications, we try to achieve the best combination of optical recognition (when the quality of the original is adequate), acquisition of electronic publications, and manual text entry.

Utilization of ready-made computer files is the principal task for creation of electronic depositories. Problems encountered here are mainly of organizational nature, as most publishers use electronic publishing systems and virtually all publications have their electronic counterparts, very often remaining unclaimed.

Another important aspect of scanning in Russian libraries is a combination of character recognition and storage of the documents' graphic patterns. In practice, it is done in the following way: first, journal content pages are scanned and used in cataloging, secondly, a journal, which is in high demand, is scanned, compressed, and stored as an electronic copy in a graphic form.

2. Speaking about software products for cyrillic text recognition (which does not exclude their ability to cope with latin texts as well), we should mention the existence of numerous products good in quality and moderate in prices. A list of Russian products carries over twenty titles, but the most popular product is FineReader 1.3 with recognition accuracy of 99.5 percent for the medium-quality originals (newspapers, type-written texts, and facsimile messages). FineReader runs under Windows and allows for:

- automatic text reading from scanner,

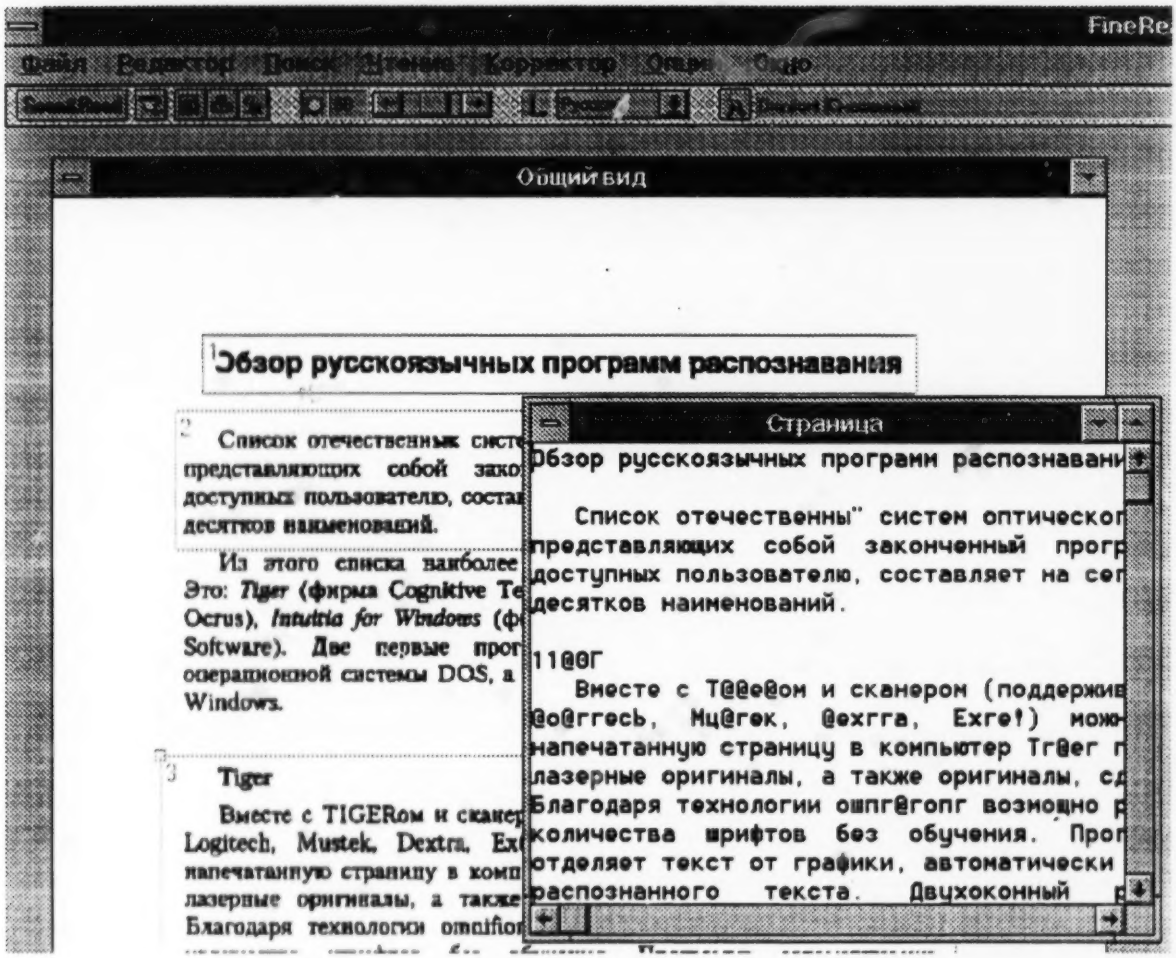
- recognition in the "intellectual" mode of vitrually any fonts without preliminary training,

- automatic separation of text and graphics,

- automatic setting of scanning brightness,

- automatic correction and spell-check,

- recognition of Cyrillic and English texts and texts with a combination of both alphabets.



The Figure shows a part of the FineReader screen.

Tiger and Author are packages running under Windows and MS-DOS. They are also well known and popular.

Cinei Form is becoming popular because it is a multi-language package working with the Russian, English, German, French, and some other European languages.

Reference

1. Sarangulian, E. Kh. Gibkaia tekhnologiia podgotovki netirazhiruemykh dokumentov s ispol'zovaniem nastol'nykh izdatel'skikh system //NTI: Seriia 1. Organizatsiia i metodika informatsionnoi raboty, N 7-8, 1991. (Flexible technology for preparation of unpublished documents by means of desktop publishing systems//STI: Series 1. Organization and methodology of information activity, No.7-8, 1991).